1. (**Finding a Large Independent Set.**) An *independent set* in an undirected graph $G = (V, E)$ is a subset of vertices $V' \subseteq V$ such that no two vertices in V' are connected by an edge of G. Recall that the problem of finding a largest independent set in G is NP-hard. In this problem, we use the Probabilistic Method to show that any graph G must contain an independent set of size at least $\frac{n}{d+1}$, where n is the number of vertices and d is the maximum degree of G. Our argument is based on the following probabilistic construction:

---

(1) assign labels $\{1, 2, \ldots, n\}$ to the vertices of G according to a random permutation.

(2) for each vertex v, if the label of v is a "local minimum" (i.e. smaller than the labels of all of its neighbors), then add v to V'.

(3) output V'.

(a) Show that the set V' output by this algorithm is indeed an independent set.

(b) Show that G must contain an independent set of size at least $\frac{n}{d+1}$. [HINT: If vertex v has degree d_v, what is the probability that v belongs to V'?]

Suppose now that we want to *derandomize* the above algorithm using the Method of Conditional Probabilities. We can proceed as follows:

---

(1) for each $i = 1, 2, \ldots, n$ in sequence, assign label i to a vertex v that maximizes the expection $\mathrm{E}[$ ... $\mid$ assignments of labels $1, 2, \ldots, i]$.

(2) output the set V' corresponding to the above label assignment, as described in the original algorithm.

(c) Fill in the blank in the Step (1) of the above algorithm. In addition, explain how to compute the expectation in Step (1).

(d) Explain briefly why the above algorithm is guaranteed to output an independent set of size at least $\frac{n}{d+1}$.

2. (**A Two-Player Game.**) MU, Exercise 6.4. [HINT: In part (b), fix a probability distribution over the removers strategies, and compute the expected number of tokens that reach position n. In particular, you will need to compute, for a fixed token, the probability that the token reaches position n. For the appropriate distribution, this quantity is (somewhat surprisingly) independent of the choosers strategy.]

3. (**Locally 2-Colorable.**) Recall that a graph (undirected, no self-loops) is 2-*colorable* if we can assign colors red and green to each vertex such that the endpoints of every edge are assigned different colors. Suppose we are told that a graph G is "locally 2-colorable", in the sense that the induced subgraph[1] on every subset of $O(\log n)$ vertices is 2-colorable. Does this imply that G itself is 2-colorable? In this problem we will see that the answer is spectacularly "no": namely, we will show that there exists a graph that is locally 2-colorable but is "very far away" from being 2-colorable, in the sense that we would have to remove a constant fraction of its edges in order to make it 2-colorable. We will prove the existence of this graph using the probabilistic method.

Throughout, set $p = 16/n$, and let G be a random graph from the model $\mathcal{G}_{n,p}$. The probabilities and expectations refer to the experiment of picking G at random.

(a) Write down the expected number of edges in G.

(b) Apply the Chernoff bound to show that with probability $1 - 2^{-\Omega(n)}$, G has at least $7(n-1)$ edges.

(c) Now fix an arbitrary assignment of colors to the vertices. Show that the expected number of violated edges (i.e., edges with endpoints of the same color) in G is at least $4(n-2)$. Deduce by a Chernoff bound that the probability there are more than $n-2$ violated edges is at least $1 - e^{-9(n-2)/8}$. [HINT: For the first part, think of the assignment of colors as being fixed *before* we choose the random edges of G. What is the value for the number of red/green vertices that minimizes the expected number of violated edges?]

(d) Show that for $n \geq 9$, with probability at least $3/4$, G is not 2-colorable even if we delete any $n-3$ of its edges. [HINT: Use the previous part and a union bound over colorings.]

(e) Show that the expected number of cycles of length exactly k in G is at most 16^k. Deduce that the expected number of cycles[2] of length at most $\frac{1}{8}\log n$ is at most $16\sqrt{n}$.

(f) Use the previous part to deduce that, with probability at least $3/4$, by deleting only $O(\sqrt{n})$ edges of G, we can obtain a graph such that the induced subgraph on any subset of $\frac{1}{8}\log n$ vertices is cycle-free (i.e., a forest – a collection of vertex-disjoint trees). (Note that a forest is always 2-colorable.)

(g) Put all of the above together to deduce that for every sufficiently large n there exists a graph $G = G_n$ on n vertices such that:

— The induced subgraph on any subset of $\frac{1}{8}\log n$ vertices of G_n is 2-colorable; and

— G_n is not 2-colorable, and remains not 2-colorable even after deleting any 0.1 fraction of its edges.

[HINT: Do be sure to take into account the fact that when we modify G to remove cycles, we may also be deleting violated edges!]

[1] An induced subgraph of a graph $G = (V, E)$ is a graph $G' = (V', E')$ where $V' \subseteq V$ and E' comprises all edges in E both of whose end-points lie in V'.

[2] Consider only cycles of length at least 3.